# Lifetime Job Mobility in the United States: A Negative Binomial Analysis

*By* Thomas Li *

*We use the negative binomial distribution to model the number of jobs individuals hold over the course of their working lives as a measure of job mobility. In particular, we seek to characterize the nature of heterogeneity in workers' propensities to change jobs while examining the time-varying dynamics and the influence of various covariates on job mobility. Our findings indicate that the negative binomial model effectively captures lifetime job counts and provides meaningful insights into mobility trends in the United States.*

A major topic of economic importance is the degree of job mobility within the labor market. Job mobility, also known as labor mobility, refers to the capacity of a worker to switch occupations or employers, with high mobility indicative of an economy with fluid redistribution of workers. On an individual level, high job mobility fosters efficient matching between employers and firms, allowing workers to find jobs that better fit their skills and career aspirations (Jovanovic, 1979) and obtain higher wages (Topel and Ward, 1992). From a macroeconomic perspective, mobility in the labor market expands national output. Indeed, recent trends in job mobility (e.g. the so-called "Great Resignation") has been a common subject of discussion among economists and Federal Reserve policy makers (Banerjee, 2022; Faccini and Melosi, 2023).

This paper seeks to understand job mobility in the United States by analyzing the number of jobs held over an individual's lifetime. Intuitively, job mobility is high if people hold many different jobs over the period of their working life, and vice versa. As this metric is non-negative and integer-valued, count models such as the negative binomial (NBD) are an appropriate tool for analysis. Employing the NBD, we seek to investigate the following questions:

1) What is the nature of the heterogeneity in lifetime job counts?
2) What are the time-varying dynamics of mobility propensity over an individual's working life?
3) What differences exist across gender, socioeconomic status, and generational group?

After describing the dataset, we present various NBD models and offer up analyses that address the questions in the above order.

* University of Pennsylvania (email: thomli@sas.upenn.edu).

## I. Data

The primary data used in this paper comes from the Bureau of Labor Statistics' 1979 cohort National Longitude Survey of Youth (NLSY79). The NLSY79 has surveyed the same cohort of Americans born between 1957–1964 every few years starting in 1979. These surveys maintain a record of the number of jobs each individual reports in each interview year as well as demographic data describing the respondent. Due to the longitudinal nature of the dataset, we have data for individuals in the Baby Boomer generation from ages 14–21 (in 1979) to ages 55–62 (in 2020)—almost their entire working lives. From the NLSY79, we further obtain information about respondents' gender, race, and socioeconomic status. The length of respondents' working lives is computed by assuming they enter the workforce the year of their first job. While this assumption is not ideal, it is the best we can do given the data. Future studies may explore different measures to test the robustness of our model results.

Our primary models of job counts (Models 1-3, 4a-b) utilize the cross-sectional sample from the NLSY79, which is representative of U.S. demographics in 1979. As such, we believe these models can generalize beyond the sample to characterize mobility for the entire Baby Boomer generation. For covariate analysis, we augment our data with the NLSY79's supplementary sample of economically disadvantaged individuals as well as the NLSY97 cohort, which mirrors the NLSY79 methodology for individuals born between 1980 and 1984. Descriptive statistics for the datasets, identified by the model they are used for, are presented in Table 1, with details on dataset differences provided in the accompanying note.

Table 1—Descriptive Statistics, by Models

| Model(s) | 1 | 2, 4a–b | 3a-d | 5, 6a-b | 7a-b | 8 |
|---|---|---|---|---|---|---|
| Respondents | 3941 | 3920 | 2455 | 2936 | 4511 | 4511 |
| % Male | 45.2 | 46.1 | 45.3 | 46.0 | 50.5 | 47.5 |
| % Female | 53.8 | 53.9 | 54.7 | 54.0 | 49.5 | 52.5 |
| % Poor | 6.6 | 6.6 | 6.1 | 50.0 | n/a | 6.4 |
| % White | 80.1 | 80.2 | 83.7 | 100.0 | 70.2 | 81.5 |
| Mean job count | 13.2 | 13.2 | 13.1 | 8.3 | 6.9 | 10.1 |

*Note:* Mean job count refers to the average number of jobs held by 2020 (Models 1-4), 1990 (Models 5-6), 2021 (Model 7), or 1998 (Model 8). "White" is defined as non-Black and non-Hispanic per the NLSY. Model 1 includes the complete NLSY79 cohort after excluding individuals with missing data or who were deceased in 2020. Models 2 and 4 use the same dataset but further exclude respondents with an unknown year of first job. Model 3 excludes individuals from the complete dataset with missing job counts before 2020 (1246 respondents) or less than 40 years of work history (240 respondents). Models 5 and 6 combine non-poor white respondents from the NLSY79 cohort with poor white respondents from the NLSY79 supplemental cohort, focusing on jobs held until 1990 due to data limitations. Model 7 is a random subset of the NLSY97 cohort, excluding individuals with missing data or who are deceased. Model 8 is similar to Model 2 but uses 1998 as the reference year for cumulative job counts.

*Source:* Bureau of Labor Statistics, National Longitudinal Survey of Youth.

## II. Lifetime Job Counts

We present two models in this section: a vanilla NBD fitted to individual-level job counts (Model 1) and a time-varying NBD that incorporates the length of an individual's working life (Model 2). In fitting a negative binomial, we assume individual-level behavior of job changes can be modeled by a Poisson process with parameter $\lambda$ where heterogeneity of $\lambda$ values across the population is given by a Gamma distribution. Table 2 presents the parameter estimates and goodness-of-fit metrics for the two models. Figures 1 and 2 respectively present histograms of the observed and expected job counts for Models 1 and 2.

Model 1 assumes all individuals operate within the same unit time—an imprecise assumption as discussed in the next paragraph. Parameter estimates for Model 1 are $r = 4.055$ and $\alpha = 0.308$. Despite using a constant unit time, the model produces a decent fit to the data as confirmed by a $\chi^2$ goodness-of-fit test ($\chi^2 = 34.5, p = 0.13$). Perhaps the distribution of observational period lengths is concentrated enough such that differences due to the length of working life are washed out. A visual analysis of Figure 1 confirms a good fit: namely, there are no obvious spikes that need to be accounted for.

Table 2—Estimated parameters and fit metrics: Models 1 and 2

| | Model 1 | Model 2 |
|---|---|---|
| $r$ | 4.055 | 4.092 |
| $\alpha$ | 0.308 | 12.966 |
| Log likelihood | −13161.6 | −13037.6 |
| $\chi^2$ goodness-of-fit | 36.517 | 36.379 |
| $p$-value (GOF) | 0.130 | 0.107 |

*Note:* The NBDs are fitted to the data using maximum likelihood. Model 1 assumes all data are in unit time whereas Model 2 accounts for different lengths of working life and employs a unit time of a year. The $\chi^2$ goodness-of-fit test statistics are respectively calculated on counts of 0–29 and 1–29 with right censoring at 30+ for both. The $\chi^2$ tests respectively use 28 and 27 degrees of freedom.

The assumption of unit time, however, fails to consider the different lengths of time individuals may be in the labor market. Differences in age and schooling, for example, may influence whether a person begins working in 1979 or 1985. Defining an individual's first year in the labor market to be the year of their first job, we find that lengths of working life vary from 14 to 55 years (mean=41.8, sd=2.5). We account for this variation by fitting a time-varying NBD. By the nature of our calculation of working life, the dataset excludes individuals who have held zero jobs. Hence, we fit a truncated time-varying NBD, as Figure 1 suggests that individuals with zero jobs exist. The estimated value of $r$ (4.055) is close to that of Model 1 and the model fit is likewise good ($\chi^2 = 36.4, p = 0.11$). As both models produce similar estimates and fits, we proceed with Model 2 due to its better story accounting for time in the labor force.

The high value of $r$ suggests greater homogeneity in people's propensities to change jobs,

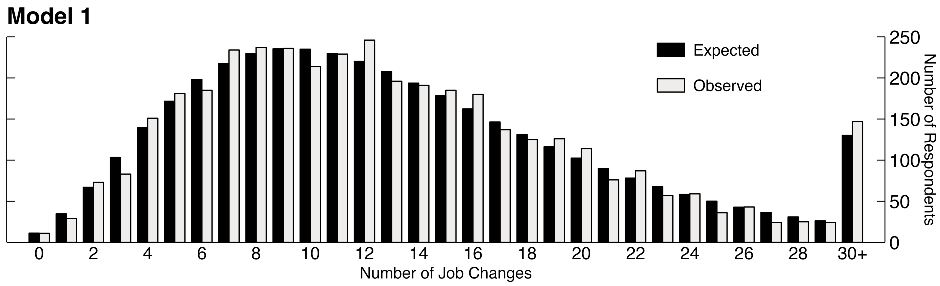


FIGURE 1. EXPECTED V.S. OBSERVED JOB COUNTS: MODEL 1

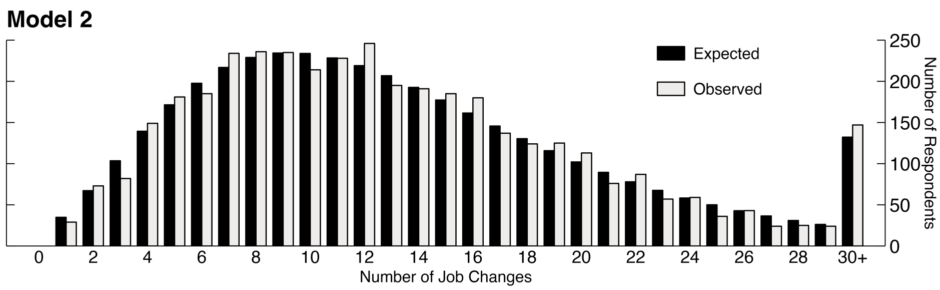


FIGURE 2. EXPECTED V.S. OBSERVED JOB COUNTS: MODEL 2

as confirmed by Figures 1 and 2. A high interior mode for $\lambda$ makes sense: most people have a handful (7–12) of jobs throughout their lifetimes with very few holding zero or one job their entire life. Some heterogeneity is nonetheless present in the spread of $\lambda$'s, aligning with previous literature on the existence of mobility heterogeneity (Mincer and Jovanovic, 1979). A Lorenz curve (Figure 3) plotting the share of job counts accounted for by each population percentile confirms the relative homogeneity of population propensities. The Gini coefficient of 0.27 suggests a low degree of concentration, with the 80$^{\text{th}}$ percentile accounting for 65% of job changes.

Assuming a typical working life of 42 years, our model estimates the average number of jobs held to be 13.25 (sd=7.5), or about a job every 3.2 years. From our truncated NBD, we also impute that 10.7 individuals (0.3%) relative to the sample size are expected to hold zero jobs after 42 years of labor eligibility—indeed, the Model 1 dataset includes 11 respondents reporting no jobs. Assuming our sample is generalizable, we would expect 0.3% of Baby Boomers to never hold a job. (These two statistics are in fact sufficient to back out parameter estimates: solving the system of equations using a mean of 13.25 and a zero-job rate of 0.3% with $t = 42$ years, we obtain $r = 4.078, \alpha = 13.004$.)

## III. Time-Varying Dynamics

Models 1 and 2 assume stationarity in individuals' mobility propensities. However, conventional wisdom suggests that younger workers are more likely than older workers to change jobs. To assess the degree to which propensities change, we present four models

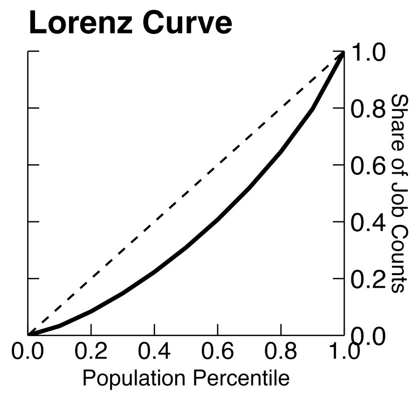


FIGURE 3. LORENZ CURVE BASED ON MODEL 2 ESTIMATES

fitted to the number of job changes in years 1–10 (Model 3a), 11–20 (Model 3b), 21–30 (Model 3c), and 31–40 (Model 3d) of a respondent's working life. Parameter estimates and goodness-of-fit metrics are presented in Table 3 and histograms of actual versus expected job counts are depicted in Figure 4.

TABLE 3—ESTIMATED PARAMETERS AND FIT METRICS: MODELS 3A–D

| | Model 3a | Model 3b | Model 3c | Model 3d |
|---|---|---|---|---|
| $r$ | 4.903 | 1.669 | 1.562 | 0.966 |
| $\alpha$ | 0.744 | 0.521 | 0.921 | 0.821 |
| Log likelihood | −3550.6 | −3001.1 | −2325.2 | −1975.1 |
| $\chi^2$ goodness-of-fit | 15.319 | 9.324 | 21.942 | 9.974 |
| $p$-value (GOF) | 0.053 | 0.316 | 0.005 | 0.267 |

*Note:* The NBDs are fitted to the data using maximum likelihood. All models use a unit time of 10 years and the $\chi^2$ goodness-of-fit test statistic is calculated on counts of 0 to 9 with right censoring at 10+. The $\chi^2$ test uses 8 degrees of freedom.

The model fits are acceptable, with 3b and 3d exhibiting high $p$-values (resp. 0.32 and 0.27) whereas 3a and 3c are less well-fitted (resp. 0.05 and 0.005). Nonetheless, a visual inspection of the histograms reveals no cause for concern. Of notable discussion is the decline in $r$ as time goes on. The population distribution of $\lambda$ becomes more heterogeneous with time as individuals' mobility propensities move toward zero. We interpret this trend to mean that individuals are less likely to switch jobs as they get older. Indeed, the histograms support the idea that older folks in the 30-40$^{\text{th}}$ year of working tend to stay put while young adults just starting out change jobs frequently. Labor economists suggest that this phenomenon can be decomposed into two effects: (1) tenure effects wherein the longer a worker stays with a firm, the less likely they are to turnover (i.e. a decreasing hazard function) and (2) pure age effects such as difficulty of relocation, search costs, and

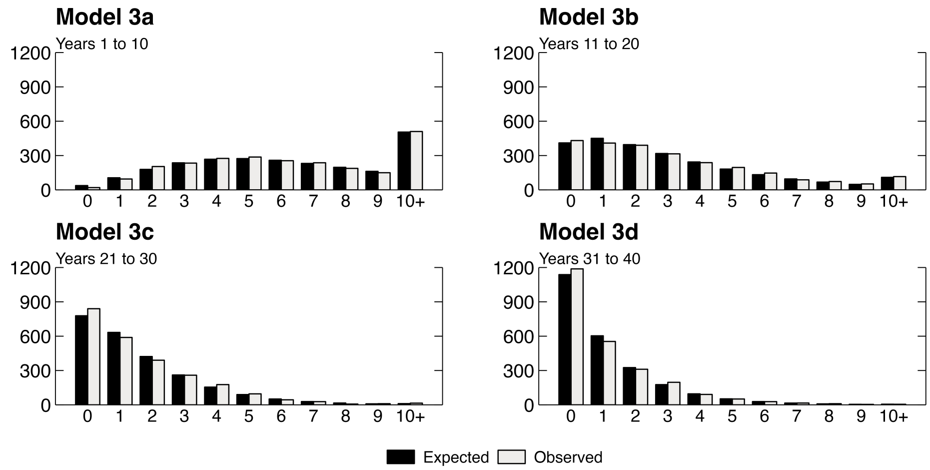


FIGURE 4. EXPECTED V.S. OBSERVED JOB COUNTS: MODELS 3A–D

*Note:* Horizontal axes display job counts. Vertical axes display number of respondents.

less desire for human capital acquisition (Mincer and Jovanovic, 1979). An interesting area for further study that is outside the scope of this paper is fitting continuous-time duration models such as the Burr XII to assess and forecast job tenure.

An immediate consequence of our findings of non-stationarity is the weakness of the models presented in this paper to project job mobility into the future. Given that $\lambda$ appears to decrease with time, any projections would overestimate larger counts and underestimate small counts. Local forecasts a few years into the future may still hold weight, but we caution against long-horizon projections.

## IV. Differences by Covariates

We now seek to characterize differences in mobility propensities due to three covariates: gender, socioeconomic status, and generational group. The motivating literature suggests that differentials in labor mobility can be explained by race (Borjas, 1984; Blau and Kahn, 1981), gender (Blau and Kahn, 1981), and wealth (Applegate and Janssen, 2020), among others. While this paper limits the analysis to male versus female, poor versus non-poor, and Baby Boomers versus Generation X, the analysis can easily be extended to other covariates.

### *A. Gender*

We fit separate time-varying truncated NBDs to female (Model 4a) and male (Model 4b) respondents, using the same dataset as Model 2 to assess the plausibility of differences between the two groups. Results are presented in Table 4 and histograms are given in Figure 5.

Similar to Models 1 and 2, the gender sub-models yield high shape parameters, with

TABLE 4—ESTIMATED PARAMETERS AND FIT METRICS: MODELS 4A-B

| | Model 4a | Model 4b | Model 2 |
|---|---|---|---|
| $r$ | 4.262 | 3.919 | 4.092 |
| $\alpha$ | 13.669 | 12.244 | 12.966 |
| Log likelihood | −6959.5 | −6076.1 | −13037.623 |
| $\chi^2$ goodness-of-fit | 24.760 | 31.114 | 36.379 |
| $p$-value (GOF) | 0.588 | 0.266 | 0.107 |
| $\chi^2$ LRT | | 4.020 | |
| $p$-value (LRT) | | 0.134 | |

*Note:* The NBDs are fitted to the data using maximum likelihood and with a unit time of a year. Due to missing zeroes, parameter estimation is computed with rescaled probabilities. The $\chi^2$ goodness-of-fit test statistics are calculated on counts of 1–29 with right censoring at 30+ and a critical threshold computed with 27 degrees of freedom. The likelihood ratio test (LRT) is conducted between Model 2 and Models 4a-b, using 2 degrees of freedom.

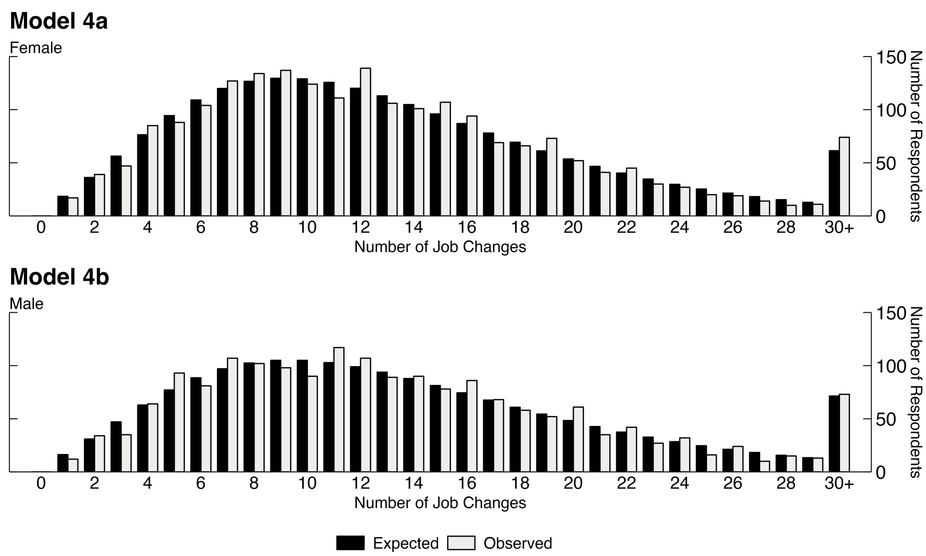


FIGURE 5. EXPECTED V.S. OBSERVED JOB COUNTS: MODELS 4A–B

$r = 4.3$ for females and $r = 3.9$ for males. We obtain good fits for Models 4a-b, as confirmed the $\chi^2$ goodness-of-fit test ($p$-values are respectively 0.59 and 0.27). The likelihood ratio test returns a non-significant $p$-value of 0.13, indicating that the improvement from considering gender is marginal. It is hard to ascertain noticeable differences in the histograms as both distributions appear similar, but the parameters provide some evidence that females tend to have higher $\lambda$'s than men due to the larger $r$ estimate. We interpret this result as suggesting that females may have a higher propensity to change jobs. This would be in line with past research suggesting that family roles may hurt women's labor force engagement and reduce their job tenures. (Blau and Kahn, 1981; Goldin and Mitchell, 2017). More data, however, is needed to confirm these results.

### *B. Socioeconomic status*

Another oft-discussed covariate is the effect of socioeconomic status on labor mobility. Using a respondent's classification in the NLSY as "economically disadvantaged" or not, we fit time-varying truncated NBDs to non-poor (Model 6a) and poor (Model 6b) respondents. Due to data limitations, these models are fit on job counts as of 1990 and only include data for non-Black and non-Hispanic respondents. For the sake of comparison, we also fit an NBD to the pooled data (Model 5). Table 5 presents the model fits and Figure 6 shows histograms of Models 6a–b.

Table 5—Estimated parameters and fit metrics: Models 5, 6a-b

| | Model 5 | Model 6a | Model 6b |
|---|---|---|---|
| $r$ | 4.610 | 4.491 | 4.772 |
| $\alpha$ | 6.649 | 6.657 | 6.700 |
| Log likelihood | −8411.7 | −4200.8 | −4207.659 |
| $\chi^2$ goodness-of-fit | 41.284 | 33.751 | 30.231 |
| $p$-value (GOF) | 0.001 | 0.009 | 0.025 |
| $\chi^2$ LRT | | 6.504 | |
| $p$-value (LRT) | | 0.039 | |

*Note:* The NBDs are fitted to the data using maximum likelihood and with a unit time of a year. Due to missing zeroes, parameter estimation is computed with rescaled probabilities. The $\chi^2$ goodness-of-fit test statistics are calculated on counts of 1–19 with right censoring at 20+ and a critical threshold computed with 17 degrees of freedom. The likelihood ratio test (LRT) is conducted between Model 5 and Models 6a-b, using 2 degrees of freedom.

The individual model fits are not great given the low $p$-values for the $\chi^2$ goodness-of-fit tests. The likelihood ratio test, however, suggests that fitting separate models on socioeconomic status improves the fit. The estimated shape parameters $r$ are on the higher end of all of our models. Our analysis of Models 3a–d suggest that this result is an artifact of using 1990 data, when workers' mobility propensities tend to be higher. It is notable that poor workers have a higher $r$ than non-poor workers, indicating higher propensities to change jobs for economically disadvantaged individuals. This finding confirms previous reports of the volatility of labor market attachment for low-wage workers (Butcher and Schanzenbach, 2018).

### *C. Generation*

Our final set of models seeks to unpack whether mobility differs by generational group. The previous models were run on data from a cohort comprised of Baby Boomers, but there is no basis to believe that mobility propensities exhibit similar features across generations.

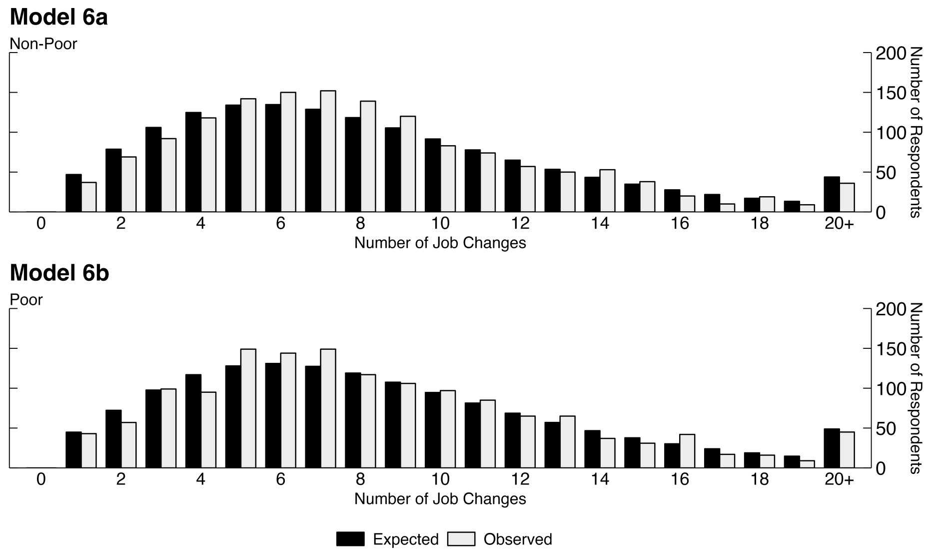


FIGURE 6. EXPECTED V.S. OBSERVED JOB COUNTS: MODELS 6A-B

To evaluate the plausibility of inter-generational stationarity, we fit a regular time-varying NBD (Model 7a) and a zero-inflated time-varying NBD (Model 7b) to job counts for Generation X using data from the NLSY97. To account for the non-stationarity of individuals' $\lambda$'s when comparing models between generations, we fit a model to data from the 1998 release of NLSY79 (Model 8). Due to differences in data availability (discussed in more detail in the Table 6 note), Models 7a-b include zeros whereas Model 8 is truncated. The model fits are presented in Table 6 and the histograms for Model 7b and Model 8 are respectively displayed in Figures 7 and 8.

Model 7a exhibits a poor fit to the data due to unexpectedly high counts of zeroes. We remedy this issue by incorporating a spike at zero—that is, we assume there is a proportion $\pi$ of hardcore-never-workers (HCNW) that do not work, leaving the remaining $1 - \pi$ of the sample to be modeled by an NBD. The existence of HCNWs is plausible. Members may include stay-at-home parents, disabled or sick people, and the ultra-wealthy (although it is interesting that HCNWs did not pop up in Model 1). The likelihood ratio test confirms that the inclusion of the spike drastically improves the model (LRT $p < 0.001$) and produces a remarkable fit to the data (GOF $p = 0.749$). We interpret $\pi = 0.025$ to mean that 2.5% of the Gen X'ers are HCNWs.

The shape parameter estimates and histograms suggest that after 20 years in the labor force, mobility propensities for Generation X tend to be smaller than that of Baby Boomers. If these data points are indicative of a larger trend, then labor mobility has been decreasing across generations, granting further evidence to economists' observations of declining job mobility since the 1980s (Azzopardi et al., 2020; Banerjee, 2022).

TABLE 6—ESTIMATED PARAMETERS AND FIT METRICS: MODELS 7A-B, 8

| | Model 7a | Model 7b | Model 8 |
|---|---|---|---|
| $r$ | 3.287 | 3.958 | 4.216 |
| $\alpha$ | 9.015 | 10.587 | 8.279 |
| $\pi$ | - | 0.025 | - |
| Log likelihood | $-12793.5$ | $-12756.6$ | $-13861.9$ |
| $\chi^2$ goodness-of-fit | 100.212 | 12.809 | 23.785 |
| $p$-value (GOF) | $< 0.001$ | 0.749 | 0.125 |
| $\chi^2$ LRT | 73.672 | | |
| $p$-value (LRT) | $< 0.001$ | | |

*Note:* All models are time-varying NBDs fitted to the data using maximum likelihood and a unit time of a year. NLSY79 does not provide the year of a respondent's first job, so we make the rough assumption that working lives start at age 20. As such, the dataset includes zeros. Model 8 is estimated in the same way as Model 2. The $\chi^2$ goodness-of-fit test statistics are calculated on counts of 0–19 (Models 7a-b) or 1-19 (Model 8) with right censoring at 20+ for both. The goodness-of-fit tests are conducted using 18 (7a) or 17 (7b, 8) degrees of freedom. The likelihood ratio test (LRT) between Model 7a and 7b uses 1 degree of freedom.

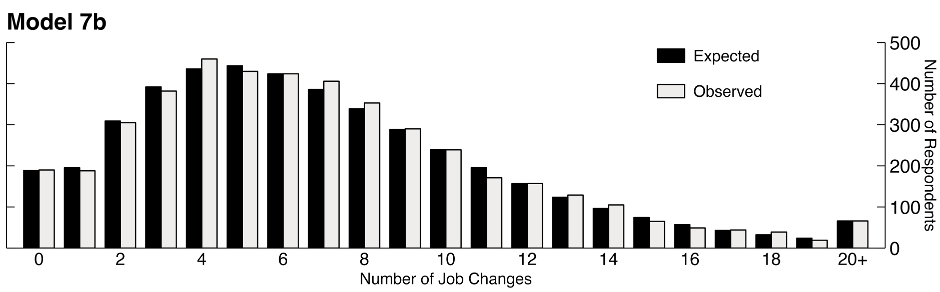


FIGURE 7. EXPECTED V.S. OBSERVED JOB COUNTS: MODEL 7B

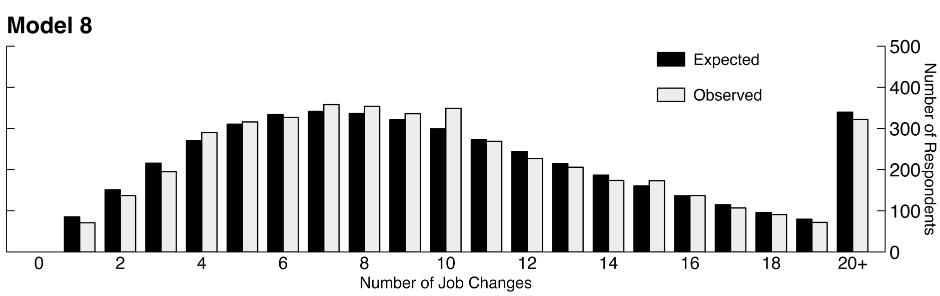


FIGURE 8. EXPECTED V.S. OBSERVED JOB COUNTS: MODEL 8

## V. Conclusion

Count data models prove to be an invaluable tool for assessing economic count data. Whereas previous research into job mobility have relied extensively on regression models, we show that negative binomial models can not only replicate results in the time-varying dynamics and covariate effects of job mobility but also yield novel insights into the degree of heterogeneity in mobility propensities. By incorporating reasonable stories of human behavior into our models, the NBD further allows us to deduce key information such as the proportion of hardcore-never-workers.

We find that job mobility propensities are relatively homogeneous and exhibit different distributions when considering gender, socioeconomic status, and generational group. A key finding across our analyses are the multitude of forces—namely, age and generation—contributing to declining labor mobility in the United States. This research opens up new avenues of exploration into the persistence and causes of these trends, as well as their broader impact on the labor market and macroeconomy.

## REFERENCES


**Applegate, Joffa, and Marco Janssen.** 2020. “Job Mobility and Wealth Inequality.” *Computational Economics*, 59(1): 1–25.

**Azzopardi, Damien, Fozan Fareed, Mikkel Hermansen, Patrick Lenain, and Douglas Sutherland.** 2020. “The decline in labour mobility in the United States: Insights from new administrative data.” *OECD Economics Department Working Papers.*

**Banerjee, Abhimanyu.** 2022. “Job Mobility During and After the Pandemic.” Federal Reserve Bank of Richmond.

**Blau, Francine D., and Lawrence M. Kahn.** 1981. “Race and Sex Differences in Quits by Young Workers.” *ILR Review*, 34(4): 563–577.

**Borjas, George J.** 1984. “Race, Turnover, and Male Earnings.” *Industrial Relations: A Journal of Economy and Society*, 23(1): 73–89.

**Bureau of Labor Statistics.** 2023. “National Longitudinal Survey of Youth 1979 cohort, 1979-2020 (rounds 1-29).” U.S. Department of Labor.

**Butcher, Kristin F, and Diane Whitmore Schanzenbach.** 2018. “Most workers in low-wage labor market work substantial hours, in volatile jobs.” *Center on Budget and Policy Priorities.*

**Faccini, Renato, and Leonardo Melosi.** 2023. “Job-to-Job Mobility and Inflation.” *The Review of Economics and Statistics*, 1–45.

**Goldin, Claudia, and Joshua Mitchell.** 2017. “The New Life Cycle of Women’s Employment: Disappearing Humps, Sagging Middles, Expanding Tops.” *Journal of Economic Perspectives*, 31(1): 161–82.

**Jovanovic, Boyan.** 1979. “Job Matching and the Theory of Turnover.” *Journal of Political Economy*, 87(5): 972–990.

**Mincer, Jacob, and Boyan Jovanovic.** 1979. “Labor Mobility and Wages.” National Bureau of Economic Research Working Paper 357.

**Topel, Robert H., and Michael P. Ward.** 1992. “Job Mobility and the Careers of Young Men*.” *The Quarterly Journal of Economics*, 107(2): 439–479.